Remote Sensing of Environment 173 (2016) 214–223

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Comparison of small-footprint discrete return and full waveform airborne lidar data for estimating multiple forest variables

Matthew J. Sumnall [a,b], Ross A. Hill [b,*], Shelley A. Hinsley [c]

[a] Virginia Polytechnic Institute and State University, Department of Forest Resources and Environmental Conservation, Blacksburg, VA 24061, USA
[b] Department of Life and Environmental Sciences, Bournemouth University, Poole, Dorset, UK
[c] Centre for Ecology and Hydrology, Wallingford, Oxfordshire, UK

ARTICLE INFO

Article history:
Received 25 February 2015
Received in revised form 17 July 2015
Accepted 27 July 2015
Available online 12 August 2015

Keywords:
Remote sensing
Forest inventory
Airborne laser scanning
Area-based regression

ABSTRACT

The quantification of forest ecosystems is important for a variety of purposes, including the assessment of wildlife habitat, nutrient cycles, timber yield and fire propagation. This research assesses the estimation of forest structure, composition and deadwood variables from small-footprint airborne lidar data, both discrete return (DR) and full waveform (FW), acquired under leaf-on and leaf-off conditions. The field site, in the New Forest, UK, includes managed plantation and ancient, semi-natural, coniferous and deciduous woodland. Point clouds were rendered from the FW data through Gaussian decomposition. An area-based regression approach (using Akaike Information Criterion analysis) was employed, separately for the DR and FW data, to model 23 field-measured forest variables. A combination of plot-level height, intensity/amplitude and echo-width variables (the latter for FW lidar only) generated from both leaf-on and leaf-off point cloud data were utilised, together with individual tree crown (ITC) metrics from rasterised leaf-on height data. Statistically significant predictive models ($p < 0.05$) were generated for all 23 forest metrics using both the DR and FW lidar datasets, with R^2 values for the best fit models in the range $R^2 = 0.43$–0.94 for the DR data and $R^2 = 0.28$–0.97 for the FW data (with normalised RMSE values being 18%–66% and 16%–48% respectively). For all but two forest metrics the difference between the NRMSE of the best performing DR and FW models was ≤7%, and there was an even split (11:12) as to which lidar dataset (DR or FW) generated the best model per forest metric. Overall, the DR data performed better at modelling structure variables, whilst the FW data performed better at modelling composition and deadwood variables. Neither showed a clear advantage at modelling variables from a particular vegetation layer (canopy, shrub or ground). Height, intensity/amplitude, and ITC-derived crown variables were shown to be important inputs across the best performing models (DR or FW), but the additional echo-width variables available from FW point data were relatively unimportant. Of perhaps greater significance to the choice between lidar data type (i.e. DR or FW) in determining the predictive power of the best performing models was the selection of leaf-on and/or leaf-off data. Of the 23 best models, 10 contained both leaf-on and leaf-off lidar variables, whilst 11 contained only leaf-on and two only leaf-off data. We therefore conclude that although FW lidar has greater vertical profile information than DR lidar, the greater complimentary information about the entire forest canopy profile that is available from both leaf-on and leaf-off data is of more benefit to forest inventory, in general, than the selection between DR or FW lidar.

1. Introduction

A forest ecosystem can be described in terms of its structural, compositional and functional properties, which can be strongly influenced by any management strategies applied to a site. The quantification of forest structure is important for a range of disciplines, as vegetation structure is related to a wide variety of ecosystem processes. However, a comprehensive understanding of the overall spatial patterns of structural variation in large forested landscapes is still largely incomplete (Anderson et al., 2008).

The management of an area is often assisted by landscape-scale monitoring (Newton et al., 2009), with a requirement of measuring both vertical and horizontal metrics. For example, the assessment of timber yields requires information on the density of trees, together with their species and size (Matthews & Mackie, 2006). Such data allow the quantification of timber yield and its associated economic value, and in addition risk assessment for fire, wind or pest damage, which are also partially dependent on canopy structure. Vertical structure is of importance in determining the species composition of ground flora (Ferris, Peace, Humphrey, & Broome, 2000), in the assessment of habitat quality for many forest-dwelling species (Hinsley, Hill, Fuller, Bellamy, & Rothery, 2009), and as an indicator of biodiversity (Ferris & Humphrey, 1999). Traditionally forest inventory data are collected

* Corresponding author.

http://dx.doi.org/10.1016/j.rse.2015.07.027
0034-4257/

through manual field observations in sample plots. The benefit of this approach can be high accuracy, but it is time consuming and expensive (Aplin, 2005).

Airborne remote sensing technologies such as lidar can characterise both horizontal and vertical structures in forested environments. The use of lidar has rapidly come to prominence in estimating forest biophysical characteristics, such as canopy height and basal area (Evans, Hudak, Faux, & Smith, 2009). Most commercial airborne lidar systems are small-footprint (i.e. <1 m) and deliver discrete return (DR) point data. The point data correspond to high intensities in the back-scattered light of the laser pulse interacting with a surface, allowing some systems to record multiple returns per laser pulse (typically 1–5). Due to limitations in the design of most multi-return airborne lidar systems, there is a sizable 'blind spot' (or dead zone) following each detected return (typically 1.2 m to 5.0 m) in which no other surfaces can be detected (Reitberger, Krzystek, & Stilla, 2008). Range resolution is determined by the length of the transmitted pulse and the maximum number of returns recorded by the sensor. The signal processing algorithms which are used to detect returns are often proprietary and differ between DR lidar sensors (Disney et al., 2010; Næsset, 2009).

Recent developments in scanning lidar technology resolve the issue of a blind spot. Small-footprint, full waveform (FW) lidar systems have become available commercially. FW lidar sensors digitise the total amount of laser energy returned to the sensor in fixed time intervals (typically 1 ns to 5 ns), providing a near continuous distribution of back-scattered laser intensity for each recorded pulse (Wagner, Hollaus, Briese, & Ducic, 2008). Instead of clouds of individual three-dimensional points, such as with DR lidar, small-footprint FW lidar devices provide connected profiles of the three dimensional scene, which contain more detailed information about the structure of the illuminated surfaces (Alexander, Tansey, Kaduk, Holland, & Tate, 2010). Each waveform consists of a series of temporal modes (or echoes), where each corresponds to an individual reflection event from an object or set of close but separated objects. Each laser pulse waveform represents complex data, which requires sophisticated processing before metrics can be generated (Chauve et al., 2009). One potential approach to derive information from the waveform is to identify proximal peaks, or returns, to present the waveform as a series of Gaussian curves; fitted by a non-linear least squares approach (Miura & Jones, 2010; Wagner, Ullrich, Ducic, Melzer, & Studnicka, 2006). The replacement of Gaussian functions with stochastic functions based on marked point processes (Mallet et al., 2010) has also been suggested as a method of processing small-footprint FW lidar data. Extracting individual returns from FW data can have the effect of removing the blind spot present in DR data that have been processed by proprietary software.

Airborne DR lidar systems have been utilised for the estimation and retrieval of various forest related variables, which are important to management and ecological monitoring. This is due to an inherent ability to provide both geo-referenced horizontal and vertical information on the structure of forest canopies, with sampling dependent on the type of lidar system used and flight configuration (Evans et al., 2009; Næsset, 2009). The most obvious vegetation measure extracted from lidar is that of canopy height. Plot- or stand-level regression analysis or non-parametric model estimates of canopy density, mean tree height, basal area and volume have been applied (Bouvier, Durrieu, Fournier, & Renaud, 2015; Hyyppä et al., 2008; Næsset, 2007). Other studies have been able to characterise understorey vegetation cover and detect suppressed trees (Estornell, Ruiz, Velazquez-Marti, & Fernandez-Sarria, 2011; Maltamo et al., 2005), assess regeneration patterns and floristic composition (Bollandsåsa, Hanssen, Marthiniussen, & Næsset, 2008; Leutner et al., 2012), and estimate deadwood volume (Kim, Yang, et al., 2009; Pesonen, Maltamo, Eerikainen, & Packalen, 2008). Lidar sensors, typically DR, can collect data at point densities sufficient to identify individual tree crowns in forest canopies and delineate crown horizontal extent and vertical depth (Kaartinen et al., 2012). Such individual tree crown (ITC) metrics have been identified as important inputs into predicative models of forest variables (e.g. Hyyppä, Kelle, Lehikoinen, & Inkinen, 2001; Persson, Holmgren, & Soderman, 2002; Popescu, Wynne, & Scrivani, 2004).

With an increasing accessibility of small-footprint FW lidar, there is a small but growing number of published studies which evaluate FW and DR lidar for the estimation of forest structural and compositional parameters. For example, Cao et al. (2014) compared statistical predictions of total living biomass obtained from DR lidar metrics (i.e. height and height variance measures, canopy return density measures, and canopy cover measures) and from FW lidar metrics (i.e. height of median energy, waveform distance, height/median ratio, number of peaks, roughness of outermost canopy, front slope angle, return waveform energy and vertical distribution ratio). They extracted the DR data by Gaussian decomposition of the FW data, and therefore the two datasets shared the same sampling rate characteristics but supplied different sets of metrics due to the way the full waveform information was processed. They found that lidar metrics related to canopy height (either DR or FW derived) were the strongest predictors of total biomass, but that there were benefits from the synergistic use of DR and FW lidar metrics in estimating the different biomass pools in the forest vertical structure. Lindberg, Olofsson, Holmgren, and Olsson (2012) outlined a method to analyse both DR and FW lidar data for the estimation of canopy vegetation volume for coniferous and deciduous forest. Estimates of volume from FW lidar were predicted more accurately than from DR lidar, especially when corrections were applied for the shielding effects of higher vegetation layers based on the Beer–Lambert Law. Allouis, Durrieu, Vega, and Couteron (2013) reported similar results where the inclusion of FW metrics improved model estimates for the prediction of above-ground biomass of individual trees, but gave slightly inferior estimates of stem volume when compared with DR lidar only. Yu, Litkey, Hyyppä, Holopainen, and Vastaranta (2014) compared DR and FW lidar for individual tree crown delineation and boreal forest species classification, reporting that FW lidar was slightly better for detecting trees, whilst DR metrics combined with FW metrics improved species classifications. Armston et al. (2013) compared DR and FW lidar data for the estimation of vertical canopy gap probability for savanna woodland, showing that models produced using FW lidar data were superior.

The use of small-footprint DR lidar data for forest inventory using an area-based regression approach is now well established (Næsset, 2007). As small-footprint FW lidar data become more readily available, early studies suggest possible benefits and potential drawbacks in moving towards these data. As yet there has been no systematic study to compare small-footprint DR and FW data for the estimation of multiple inventory variables from across a forest profile. This study addresses this research gap, comparing point cloud data and derived products from DR lidar and from Gaussian decomposition of FW lidar. The work of Cao et al. (2014) compared standard DR height metrics with newer sets of FW lidar metrics, and specifically avoided investigating the effects of higher density point clouds provided by FW lidar decomposition. Here we specifically focus on a comparison between the different information content on forest vertical and horizontal structure and recorded return pulse characteristics in DR and FW-derived point clouds. We assess 23 common forest inventory variables covering all forest vegetation layers (canopy, shrub and ground layer) and both living and dead wood. Airborne DR and FW lidar data were acquired simultaneously under both leaf-on and leaf-off conditions, and variables from both (including point cloud and ITC-derived lidar variables) are used in area-based regression modelling of forest inventory variables. The wider context of this work was forest condition assessment.

2. Data and methods

2.1. Study site

The study site is located within the New Forest National Park, between Southampton and Bournemouth, in southern England (lat: 50°